Lessons of Building a Centralized Land Use Information System For East Boston: Microcomputer GIS Implementation Using PC ARC/INFO

Jeffrey Brown

Bizhan Azad

Boston Redevelopment Authority
Policy Development and Research Department

August 9, 1989

Revised November 17, 1989

Presented at the 1989 URISA Conference, Boston, MA

Jeffrey Brown
Bizhan Azad
Department of Policy Development and Research
Boston Redevelopment Authority
Charlestown Navy Yard, Building # 33
Charlestown, MA 02129

LESSONS OF BUILDING A CENTRALIZED LAND USE INFORMATION SYSTEM FOR EAST BOSTON: MICROCOMPUTER GIS IMPLEMENTATION

Abstract: Boston's planning agency designed and implemented a microcomputer-based land use information system to put more and better data at the fingertips of planners and community zoning advisory committee members. The primary tasks for the project were: (1) to pull together and integrate parcel-level data in electronic form from various city departments; (2) to link these data to boundary/location information (in electronic form) about each parcel for graphic representation and visual display; and (3) to implement a dozen standard easy-to-do queries of the system based on most frequently asked questions by BRA planners and the community members. In short, the aim was to implement a geographic information system (GIS) on a desktop microcomputer. The software used was PC ARC/INFO.

This paper describes the stages of the project and documents the issues and concerns that appeared and how they were tackled. The emphasis of the paper is on learning concrete lessons that are useful for implementing desktop GIS. Microcomputer-based GIS tools present challenges and opportunities in an operational environment like a planning agency where timely information is needed for specific tasks. For urban planners and analysts, this paper offers an operational evaluation and useful tips on implementation of microcomputer-based GIS.

INTRODUCTION

What if each of you returned to your office or classroom on Friday to find an assignment: Implement a geographic information system on a microcomputer in the next twelve months? Here's a budget, here's your staff, go.

Between mid-1988 and mid-1989, the Policy Development and Research Department of the Boston Redevelopment Authority (BRA) did get an opportunity to build a GIS for a section of Boston. A reasonable budget, an expectant community advisory committee, and a motivated staff provided the necessary thrust for a successful implementation.

First, this paper briefly describes the project. Then, the authors recommend a way of thinking about building a GIS, comment on the limits of a GIS software in a microcomputer DOS environment, and offer a few tips on implementation.

BACKGROUND

The BRA is the planning agency for the City of Boston, with responsibilities that include neighborhood and harbor planning, zoning, review of private development projects, and research and analysis of the city's economic, demographic, and land use characteristics.

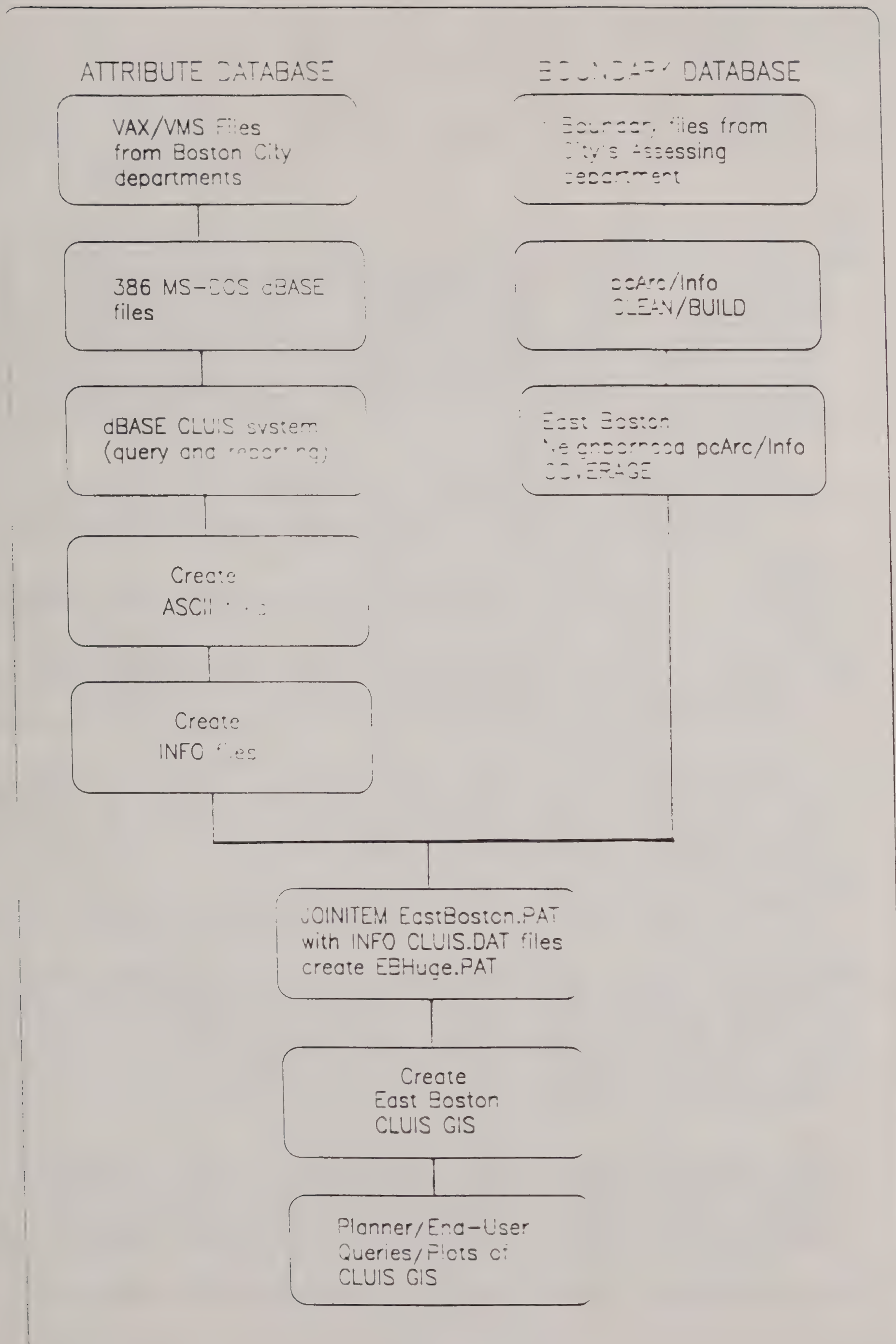
In the East Boston planning district, a community group, Air Impact Relief, Inc., won a lawsuit against the Massachusetts Port Authority (Massport) over issues regarding Logan Airport expansion in East Boston. As part of the settlement, Massport agreed to hire the BRA to design and implement a land use information system for East Boston. In this case, the community defined the problem--airport impact on land use in the neighborhood--and identified the need for a centralized system of land use data and access.

When the BRA began the Centralized Land Use Information Study (CLUIS) in July, 1988, access to land use data was cumbersome for planners and the neighborhood planning and zoning advisory committees. Needed information was available on mainframe computers in multiple files, in varying formats, from multiple city departments. The mainframe systems were fine for their designed purposes, such as assessing property, issuing tax bills, and recording building permit applications. However, research on a particular parcel of land was time-consuming, and compilation and analysis of data on multiple parcels was tedious at best. To complicate matters, the BRA was operating off-site with a mini VAX and numerous microcomputers, without a link to City Hall's mainframe computers.

The CLUIS project presented a challenge to the BRA to produce useful products for planning and zoning within a 12-month period while changing the operational environment to accommodate the information system. The project also presented timely opportunities for analysts and planners in the midst of BRA efforts to enhance the quality of life in neighborhoods through rezoning and planning in the context of a strong local economy.

The basic approach of CLUIS was to (1) extract several data files from city departments into a common format on a City Hall VAX, (2) transfer tapes to the BRA mini-VAX, then (3) download the files to a BRA microcomputer, (4) link the files to a boundary/location file, (5) analyze the information using a desk-top GIS, and (6) produce reports and thematic maps to answer queries from a community advisory committee and from planners. See chart of project elements. Interactive changes to

ELEMENTS OF CLUIS GIS PROJECT



the database were not included in the database design. See list of files in Appendix 1.

The hardware used on the project included a 386 microcomputer, a Dell 310, with a 322-megabyte hard disk, three megabytes of extended memory, a [illegible]-megahertz CPU, one 5.25-inch disk drive, and a back-up tape drive. Peripherals included a math co-processor, an 14-inch color monitor, a VGA card, a mouse, and a Hewlett-Packard Draftmaster 1 E-size plotter. The software included PC ARC/INFO Version 3.3, dBASEIII Plus, and device drivers and text editors. A full list of hardware and software is provided in Appendix 2.

In addition to BRA staff, consultants Bowen & Hayes, Inc. carried out much of the database design work, and Simon Lewis from the MIT Computer Resources Lab provided timely technical assistance in the GIS development. The project produced 48-inch by 36-inch plots of the parcels with color shading for various queries. For example, the plot below presented parcels that have off-street parking spaces, with shading to represent ranges for the number of spaces. The staff prepared screen demonstrations of polygon shading as well as menu-driven access to the attribute data base. Initial plots and access were designed to respond to the most commonly asked questions from the community review committee.

However, the focus of this paper is not products or a chronology of the system design and development. The authors would like to share some of the lessons learned that others may find useful in implementation of a GIS on a microcomputer.

THINKING ABOUT YOUR GEOGRAPHIC INFORMATION SYSTEM

Imagine that you will start building a GIS next week. General knowledge of and theoretical understanding of geographic information systems will help you choose a GIS software package that will be appropriate for your project. But progress toward implementation of a GIS comes with specific knowledge of your software package.

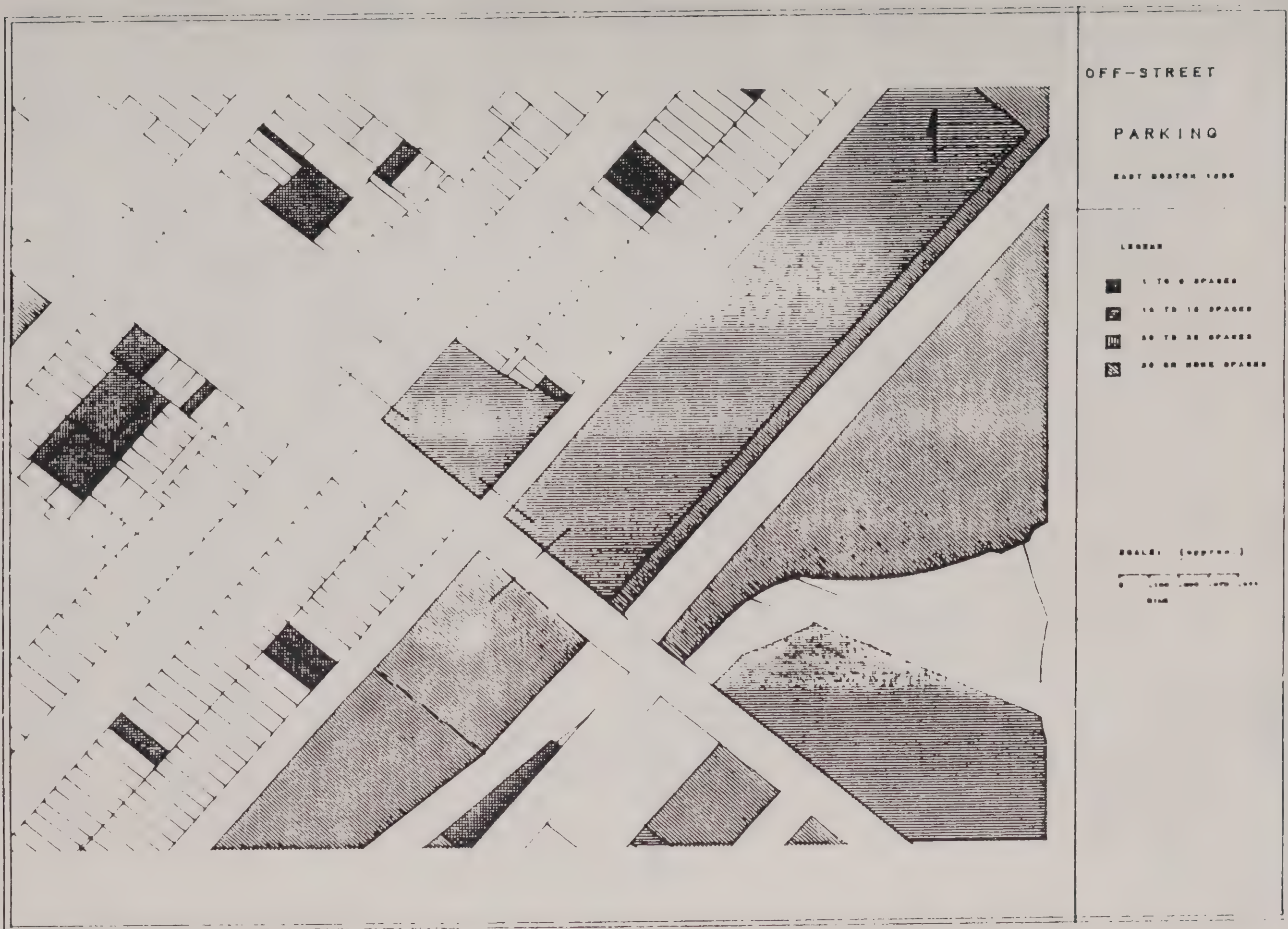
Think of your favorite violinist, say Isaac Stern, and his fine but imperfect instrument. He knows his violin. He may be ignorant of the molecular structure of the wood or the physics of the sound, but he knows intimately the strings, neck, and bow, and how they work together to produce exact pitches and the subtlest nuances.

Know your GIS package. How? Expect problems; find problems; solve problems; document the problems; push the package to its limits, especially on the PC.

Expect problems. Unless you have enough stamina and retention to read and reread the user manuals cover to cover, you will run into problems. Expect days when you make no progress because you have found yet another problem. As good as the PC ARC/INFO manuals are, the writers cannot spell out everything. Just as the violinist cannot learn from a musical score how to play a difficult passage, you must use the software to find the best way to build your GIS. Play wrong notes; get your fingers tangled.

Find problems. The staff spent many days finding technical problems in our GIS. Those problems were not bugs in the PC ARC/INFO 3.3 software, but characteristics of the software that we had to understand in order to build and label polygons, build and attach information to each polygon, select geographic features, and query the system.

Solve problems. Solving problems is satisfying. Our primary problem-solving tools



OFF-STREET
PARKING
EAST BOSTON
SCALE: (approx.)

were the PC ARC/INFO user manuals which are well written and informative. The vendor's technical support staff provided more detail where we could not read between the lines. We tried to solve as many problems as possible, even as the project deadline approached.

Document problems. We learned the importance of documentation of problems. We kept a log book where we described problems and the sequence of commands for editing, plotting, and data manipulation. Those notes became critical when we discovered that we had added a duplicate label to a street polygon. We were able to quickly correct our mistake from two months earlier because we had notes on the location and label of the particular street polygon. In another instance, we had to solve a problem a second time because we had failed to enter the first solution in the log. Perhaps the most important lesson of this project is: take a few extra seconds as you go to write down the problem, your approach, and details of the solution.

Push the software to its limits. Practical use of the software makes some limits obvious, but an effort to push the system will inform your thinking and planning for future GIS applications. This paper might have been written as "Understanding the Limits of Microcomputer GIS Software in the DOS Environment." In that regard, we offer comments on seven of the limits and the lessons learned from them.

LIMITS

1. The boundary location file. The City of Boston's Department of Management Information Services provided a boundary file for East Boston that defined the parcel boundaries. The file was digitized in 1988 in state plane coordinates of 8 digits. However, when we looked at the file in the editing module we found several errant lines extending from points within the East Boston map. This may have been an integer overload in the system. Time constraints called for a quick solution, not a verification of the exact problem.

We trimmed two digits from the left of each coordinate to produce 6-digit coordinate. This step would not alter the accuracy of the lines on the ground, but would raise problems if we ventured far enough from Boston where the initial two digits would make a difference. The lines formed by the 6-digit coordinates looked clean and accurate compared to the corresponding lines on planimetric maps.

2. Creation of Polygons. The command in the software that cleans dangling arcs (lines) and creates polygons from the arcs will not work if it finds a polygon with more than 5,000 arcs. That statement in the manual gave us no worries, but in practice, the software finds complex polygons formed by streets. East Boston, with 7,000 parcels in six square miles, was composed of 21,000 arcs, but the software was finding more than a quarter of the arcs in a single polygon.

With our coverage, the polygon creation (topology) function took about 50 minutes, making trial and error in adding arcs rather time-consuming. Eventually, we solved the problem by adding arcs to close off streets as shown on the plot of a section of East Boston. By giving the arcs identification numbers in a series higher than any of the boundary file arcs, we were able to later, for final plotting, assign the background color to those added arcs.

3. Loading data into the GIS database manager. This step turned out to be quite time-consuming. Our initial data management was carried out with a user-friendly, fast database manager developed for micro-computer applications. Thirteen files

ARCS ZOOM

EAST BOSTON

EBL 2

together amounted to about 1,500 bytes of information. Our first attempt to load 7,000 records, 400 bytes long, into the GIS (using INFO) took four hours to run. We then found an alternate method using a command (TABLES) in another module, which took much less time. However, that method has a limit of 132 bytes, which required allocation of data into about 25 files prior to loading.

4. Building polygon attribute tables. In the PC ARC/INFO system, characteristics of polygons, in this case land parcels, are brought together in a polygon attribute table (PAT). We selected 1,000 bytes of information to add to the PAT. As the PAT grows, adding items of data take many hours. For example, when we added 100 bytes to a PAT with records of 800 bytes, the operation ran for 25 hours. Also, the joining function is not totally relational. That is, the function takes only the first occurrence of an item in a given parcel. Consequently, multiple occurrences are not available for analysis or display in the GIS without multiple polygon attribute tables.

5. Maintaining data integrity. First, in our data base design, planners would have data files but would not be able to edit the data. We made one person responsible for updating records, to be done in the micro-computer database manager.

Second, any editing of the coverage requires care to avoid misalignment (or improper indexing) between the attributes and the polygons. In PC ARC/INFO, records are identified by both a record number and a coverage identification number (parcel in this case). We sorted on the coverage ID in one operation to speed processing, but failed to sort on record number for a subsequent operation. Results included an airport-related business landing on a colleague's house lot.

As the GIS files become larger and more complex, maintenance of data integrity requires attention and coordination.

6. Performance of the system. The software took 5 to 6 minutes to draw all 7,000 polygons in East Boston. This seemed especially slow when we started demonstrating the system to other staff members. Another module in the software offers a function that we used to dissolve the parcels lines to leave block lines. See plot of all blocks in East Boston. The system could draw the 230 blocks in less than one minute, allowing us to proceed with selection of parcels based on land use, sales, or whatever attributes we desired. The software allows for color shading of parcels over the block outline. As the coverage grows, operations slow down.

We strongly recommend as much expenditure on computer hardware as possible. Our system was adequate, but too often the system was slow enough to prevent the solving of a particular problem in one day. Much time was lost retracing steps days later when we returned to the problem. Soup up your microcomputer!

7. Set-up of the computer. The space and memory demands of the GIS, in addition to the requirements of the data base files and related software, called for a multiple-use set-up for the 386 computer. Two configurations separated access to the dBASE files and programs from the GIS files. The hard disk was divided into ten, 32-megabyte drives. We found that we needed to keep one drive empty in order to have enough space to create polygons. That function may take up 3 to 14 times the amount of space occupied by the coverage to be cleaned due to use of temporary scratch files.

EAST BOSTON

EBL2WPB1

SCALE: (approx.)

LESSONS LEARNED

These limits might be summarized in seven lessons learned on this project.

1. Regarding data such as the boundary file that come from other departments, maintain good relations with your sources of data because you will probably need a second transmittal.

2. Before asking the GIS software to create polygons, get a plot of the parcel lines, and look for street networks that could form large polygons. Then try the operation to better define the problem.

3. Expect to spend many hours loading data into the GIS initially and in updates.

4. Expect design problems such as our multiple occurrences example, and be prepared to spend time solving them.

5. Think through the sorting of files, and after sorting, check known parcels on a plot to quickly identify problems.

6. Purchase as much high-speed disk space and memory as possible for your microcomputer to improve performance and problem-solving.

7. Either obtain an extra computer with plenty of disk space or take time to regularly clean up the directories and leave a free drive for running the space-intensive functions.

The eventual users of a GIS, planners and neighborhood advisory committees in our case, do not need to know the software inside out. Our goal is to establish an accessible desk-top GIS that features standard queries, query by example, and standard printouts and plots. Our East Boston planner and community advisors are just beginning to use our prototype GIS.

As an implementor of a GIS, the questions that you should ask yourself throughout the process are: How will I use this lesson, this knowledge to reproduce the system later for another geography, another set of attributes? How could I more efficiently implement a system? Am I documenting the process in enough detail to reproduce the results?

One other item that could be a novel by itself: management and administration can be incredibly time-consuming for a project that involves many people in many departments and agencies in non-routine tasks.

USE AND FURTHER DEVELOPMENT OF THE SYSTEM

The authors look forward to the next stages of the land use information system. The community planning and zoning advisory committee and the East Boston planner have begun to use the system for queries about particular properties or about neighborhood areas adjacent to the airport property. The screen access to the data base through menus is the easiest part of the system to use; computer-literate community members may want to sit at the keyboard themselves. The mapping component of the system is not user friendly, necessitating operation by a staff person.

In this case, the community will likely use the system to generate alternative

rezoning plans with special attention to areas on the airport periphery. Most likely, planning will tend to discourage traffic-generating development.

In other neighborhoods of Boston, the planning emphasis will be on development potential. The current land use information system could be expanded to cover any or all of Boston's 21 other wards. The data are readily available in usable form with the exception of zoning data which are not automated. The BRA could use land use information from an expanded GIS for planning efforts in neighborhoods such as Roxbury where community advisors seek more development and economic opportunity. Build-out analysis with graphic locations would be a valuable tool in those redevelopment efforts.

The land use information system, whatever the geographic coverage, needs fresh information periodically or it loses its value as a tool for analysis and planning. The BRA and other city departments will next set up a process whereby departments with source data supply updated information, with corrections discovered through use of the GIS, at agreed upon intervals.

The BRA's prototype system for East Boston demonstrates that the use and further development of the centralized land use information system could be carried out on microcomputers. However, the system for East Boston, about one-twentieth of the city, pushed against the limits of the hardware. The BRA needs to explore hardware options for an expanded system. Concurrently, the BRA and the City need to coordinate the current development of a parcel-based information system on the City's Intergraph system with expansion of the microcomputer GIS. While the Intergraph system has the requisite power and capacity to handle a city-wide GIS, less intensive users such as planners and community advisors will be best served by access to a land use information system on a personal computer.

APPENDIX 1. LIST OF DATA FILES IN EAST BOSTON CENTRALIZED LAND USE INFORMATION SYSTEM (CLUIS)

File Name	Description
Field Inspection	Assessor's parcel file, FY88, with changes and additions from field inspection in August, 1988.
Zoning Board of Appeals	BRA file of zoning petitions, 1985
Addressee	Person designated to receive taxbills from the City Assessing Department. If no addressee, ma goes to owner.
Commercial	Structure and occupancy data for each commerci or exempt parcel from Assessing Department.
Fires	Building fires 1986-88, from Fire Department.
Inspectional Services	Information on all building permits from Inspectional Services Department, 1985-88.
Mail	Current mail address used by Assessing for taxbills (goes with Addressee file).
Parcel	Current Assessing data on each parcel, including land use, size, ownership, location.
Rent Equity	Rent controlled and vacancy decontrolled units from the Rent Equity Board file.
Residential	Current Assessing data on structure, occupancy, parking.
Sales	Sales date, price, grantee, grantor, 1983-1988 from Assessing Department.
Tax History	Includes total tax owed by fiscal year 1985-89, from Assessing Department.
Tax Title	Accumulated Balance of back taxes owed, from Assessing.

APPENDIX 2. CLUIS HARDWARE AND SOFTWARE

Hardware:

Dell 386 computer (model 310)
- 20 megahertz CPU;
- 322-megabyte hard disk with 18 milliseconds access;
- one 1.2-megabyte, 5.25-inch disk drive;
- one back-up tape drive;
- extended memory, 3 megabytes;
- math co-processor: 80-387;
- 14-inch color monitor: NEC Multisync 2A;
- Video Graphic Array (VGA) card;
- Printer: PC's Limited Printer System 800, wide carriage, dot matrix;

IBM PS/2 Model 80:
- 44-megabyte hard [illegible]
- color display 8514;
- Houston Instruments HI-PAD Plus digitizer 12 X 12;
- Printer: Hewlett Packard Laserjet Series II.

Hewlett Packard A-size desk plotter;

Hewlett Packard Draftmaster [illegible] size plotter;

Software:

PC ARC/INFO, version 3.3, geographic information system;

dBASE III Plus, data base manager;

Graphic Software Systems CGI Device Drivers;

Epsilon text editor;

AWK string manipulation language.

Approximate disk space required for CLUIS:

- Coverage of East Boston in pcARC/INFO -- 6 megabytes;
- INFO files in PC ARC/INFO -- 20 megabytes;
- Data files for East Boston -- 20 megabytes;
- Software PC ARC/INFO and others -- 20 megabytes.